A customized non-exclusive clustering algorithm for news recommendation systems

ABSTRACT

Clustering is one of the main tasks in machine learning and data mining and is being utilized in many applications including news recommendation systems. In this paper, we propose a new non-exclusive clustering algorithm named Ordered Clustering (OC) with the aim is to increase the accuracy of news recommendation for online users. The basis of OC is a new initialization technique that groups news items into clusters based on the highest similarities between news items to accommodate news nature in which a news item can belong to different categories. Hence, in OC, multiple memberships in clusters are allowed. An experiment is carried out using a real dataset which is collected from the news websites. The experimental results demonstrated that the OC outperforms the k-means algorithm with respect to Precision, Recall, and F1-Score.

Keyword: Clustering algorithm; Non-exclusive clustering; News recommendation; Similarity weight